

**PROBABILISTIC FORECASTS**  
**FROM THE NATIONAL DIGITAL FORECAST DATABASE**

By

Roman Krzysztofowicz and W. Britt Evans

University of Virginia  
Charlottesville, Virginia

March 2007

---

*Corresponding author address:* Professor Roman Krzysztofowicz, University of Virginia, P.O.  
Box 400747, Charlottesville, VA 22904-4747. E-mail: [rk@virginia.edu](mailto:rk@virginia.edu)

## **ABSTRACT**

The Bayesian Processor of Forecast (BPF) is developed for a continuous predictand. Its purpose is to process a deterministic forecast (a point estimate of the predictand) into a probabilistic forecast (a distribution function, a density function, and a quantile function). The quantification of uncertainty is accomplished by extracting and fusing two kinds of information from two different sources: (i) a long sample of the predictand from the National Climatic Data Center, and (ii) a short sample of the official National Weather Service forecast from the National Digital Forecast Database. The official forecast is deterministic and hence deficient: it contains no information about uncertainty. The BPF remedies this deficiency by outputting the complete and well-calibrated characterization of uncertainty needed by decision makers and information providers. The BPF comes furnished with (i) the meta-Gaussian model, which fits well meteorological data as it allows all forms of marginal distribution functions, and non-linear and heteroscedastic dependence structures, and (ii) the statistical procedures for estimation of parameters from asymmetric samples and for coping with nonstationarities in the predictand and the forecast due to the annual cycle and the lead time. A comprehensive illustration of the BPF is reported for forecasts of the daily maximum temperature issued with lead times of 1, 4, and 7 days for three stations in two seasons (cool and warm).

## TABLE OF CONTENTS

<b>ABSTRACT</b> .....	ii
<b>1. INTRODUCTION</b> .....	1
1.1 The Uncertainty Quantification Problem .....	1
1.2 Bayesian Processor of Forecast .....	1
1.3 Information Fusion .....	2
1.4 Modeling Approach .....	3
<b>2. BAYESIAN PROCESSOR</b> .....	4
2.1 Concept .....	4
2.2 Characterization of Uncertainties .....	4
2.3 Bayesian Revision .....	5
2.4 Bayesian Meta-Gaussian Model .....	6
2.5 Modeling and Estimation .....	6
<b>3. THE PRIOR DISTRIBUTION FUNCTION</b> .....	7
3.1 Climatic Sample .....	7
3.2 Standardization .....	7
3.3 Empirical Distribution Function .....	8
3.4 Parametric Distribution Function .....	9
3.5 Stationary Prior Distribution Function .....	10
3.6 De-standardization .....	10
3.7 Comparison of Models .....	12
<b>4. THE LIKELIHOOD FUNCTION</b> .....	13
4.1 Joint Sample .....	13
4.2 Adaptive Scheme .....	14
4.3 Joint Standardization .....	14
4.4 Marginal Distribution Function .....	16
4.5 Meta-Gaussian Likelihood Model .....	17
4.6 Likelihood Parameters .....	18
4.7 Forecast Informativeness .....	19
<b>5. PROBABILISTIC FORECAST</b> .....	21
5.1 Posterior Parameters .....	21
5.2 Forecasting Equations .....	21
5.3 Posterior Functions .....	22
5.4 Limit of Predictability .....	23

<b>6. DISCUSSION</b>	24
6.1 Forms of Heteroscedasticity	24
6.2 Revision of Uncertainty	24
6.3 Comparison of Performance	25
6.4 Generalization of Parameters	25
<b>7. CLOSURE</b>	26
<b>ACKNOWLEDGMENTS</b>	27
<b>APPENDIX A: WEIBULL DISTRIBUTION</b>	27
<b>REFERENCES</b>	28
<b>TABLES</b>	30
<b>FIGURES</b>	35

# **1. INTRODUCTION**

## **1.1 The Uncertainty Quantification Problem**

The National Digital Forecast Database (NDFD) was designed by the National Weather Service (NWS) to store the official forecasts of the sensible weather elements produced by the NWS field offices throughout the United States (Glahn and Ruth, 2003). The official forecasts are subjective in that they are made judgmentally by human forecasters with the support of software systems and based on information from multiple sources, including output from numerical weather prediction models and guidance from the national centers. With the exception of the occurrence of precipitation, which is forecasted in terms of probability, all other weather elements are forecasted deterministically. Hence the deficiency of the NDFD: it contains no information about forecast uncertainty (Ryan, 2003).

To remedy this deficiency, the Meteorological Development Laboratory of the NWS began developing statistical techniques for assessing the uncertainty in forecasts disseminated through the NDFD (Peroutka et al., 2005). This article presents a solution to the same problem, but via a different technique and in a different format.

## **1.2 Bayesian Processor of Forecast**

The Bayesian Processor of Forecast (BPF) for the NDFD is a specialized application of the Bayesian theory of probabilistic forecasting formulated and tested in various settings over the past two decades (e.g., Krzysztofowicz, 1983; Alexandridis and Krzysztofowicz, 1985; Krzysztofowicz and Watada, 1986; Krzysztofowicz and Reese, 1991; Krzysztofowicz, 1999; Krzysztofowicz and Kelly, 2000a).

The BPF developed and illustrated herein quantifies the uncertainty in a deterministic forecast of the daily maximum temperature — one of the predictands selected by the NWS for development

and testing of their technique (Peroutka et al., 2005). In general, this BPF is applicable to any continuous predictand. The inputs to and the outputs from the BPF are as follows (Fig. 1). In the estimation phase, the inputs are a climatic sample of the predictand, and a joint sample of the forecast and the predictand for a given forecast point (a grid point or a station) and a specified lead time; the outputs are the values of parameters (as the BPF is entirely parametric). In the forecasting phase, the input is a *deterministic forecast* (a point estimate of the predictand) and the output is a *probabilistic forecast* (a distribution function, a density function, and a quantile function). Thus, the BPF outputs the complete and well-calibrated characterization of uncertainty needed by rational decision makers who use formal decision models and by information providers who want to extract various forecast products for their customers (e.g., quantiles with specified exceedance probabilities, credible intervals with specified inclusion probabilities, probabilities of exceedance for specified thresholds).

### 1.3 Information Fusion

In concept, the BPF quantifies the *total uncertainty* about a predictand, given a deterministic forecast. The quantification of uncertainty is accomplished via Bayes theorem, which extracts and fuses two kinds of information from two different sources (Fig. 1): (i) Information about the natural variability of the predictand is extracted from a climatic sample, which may be retrieved from the National Climatic Data Center (NCDC). (ii) Information about the predictive performance of the deterministic forecast is extracted from a joint sample, which may be retrieved from the NDFD.

The size of each sample may be limited, not only by the length of the record, but also by the requirement of statistical homogeneity. For instance, the nonstationarity of climate may necessitate a truncation of the available climatic sample; and a significant modification in the forecasting system (e.g., hiring of an experienced or a novice forecaster by a field office; an improvement of

a numerical weather prediction model by a national center) may necessitate a truncation of the available joint sample. However, the most important fact here is that the joint sample is typically much shorter than the climatic sample. This gives rise to the statistical problem of information fusion, which can be solved correctly only by a proper application of Bayes theorem.

#### **1.4 Modeling Approach**

The key challenges that must be overcome during the development of a proper technique for quantifying uncertainty include the non-stationary behavior of the meteorological time series due to seasonality, the non-Gaussian form of the probability distributions of meteorological variates, and the non-linear and heteroscedastic dependence structure between the forecast and the predictand. These features of meteorological data, and the ways the BPF handles them, are explained and illustrated throughout the article.

The article is organized as follows. Section 2 outlines the theoretic foundation of the BPF, and defines its major components. Section 3 details the modeling and estimation of the first component: the prior (climatic) distribution function. Section 4 does the same for the second component: the family of the likelihood functions. Section 5 presents examples of probabilistic forecasts. Section 6 discusses several attributes of the BPF and the empirical results obtained for three stations. Section 7 summarizes the unique advantages of the BPF.

## 2. BAYESIAN PROCESSOR

### 2.1 Concept

Let  $W$  be the *predictand* — a continuous variate whose realization  $w$  is being forecasted. Let  $X$  be the *estimator* — a continuous variate whose realization  $x$  constitutes a point estimate of  $W$ .

From the viewpoint of the NWS,  $x$  is the official, deterministic forecast of  $W$  prepared by a field office. From the viewpoint of a rational decision maker, who recognizes the uncertainty about  $W$ , forecast  $x$  is merely a realization of the predictor  $X$  — a piece of information that may reduce the uncertainty about  $W$ , but cannot eliminate it. What the rational decision maker then needs is not a number  $x$ , but a function of  $w$  — the distribution function of predictand  $W$ , conditional on the predictor realization  $X = x$ . The purpose of the BPF is to supply such a conditional distribution function.

### 2.2 Characterization of Uncertainties

The inputs to the BPF are the prior density function and the family of likelihood functions. These inputs are defined and interpreted as follows.

Let  $g$  denote the *prior density function* of the predictand  $W$ . This density function characterizes the natural variability of  $W$ . Equivalently, it characterizes the uncertainty about the predictand that exists before the NWS issues a forecast. This uncertainty may be called the climatic uncertainty (or the prior uncertainty) and may be quantified based on climatic data.

Let  $f(\cdot|w)$  denote the density function of the predictor  $X$ , conditional on the hypothesis that the realization of the predictand is  $W = w$ . This density function characterizes the variability of  $X$  on all those occasions on which  $W = w$  is observed. What is needed is a family of the conditional density functions  $\{f(\cdot|w) : \text{all } w\}$ . Then, for a fixed forecast  $X = x$ , a function  $f(x|\cdot)$  exists; it is called the likelihood function of  $W$ . More generally, there exists a *family of the*



*likelihood functions*  $\{f(x|\cdot) : \text{all } x\}$ . The family  $f$  quantifies the stochastic dependence between the predictor  $X$  and the predictand  $W$ .

### 2.3 Bayesian Revision

The Bayesian procedure for information fusion and revision of uncertainty involves two steps. First, the *expected density function*  $\kappa$  of predictor  $X$  is derived via the total probability law:

$$\kappa(x) = \int_{-\infty}^{\infty} f(x|w)g(w) dw. \quad (1)$$

Second, the *posterior density function*  $\phi(\cdot|x)$  of predictand  $W$ , conditional on a deterministic forecast  $X = x$ , is derived via Bayes theorem:

$$\phi(w|x) = \frac{f(x|w)g(w)}{\kappa(x)}. \quad (2)$$

In concept, Bayes theorem revises the prior density function  $g$ , which characterizes the climatic uncertainty about  $W$ , given forecast  $X = x$ . The extent of the revision is determined by the likelihood function  $f(x|\cdot)$ , which characterizes the degree to which  $X = x$  reduces the uncertainty about  $W$ . The result of this revision is the posterior density function  $\phi(\cdot|x)$ ; it quantifies the uncertainty about  $W$  which remains after the NWS issues forecast  $X = x$ .

The corresponding *posterior distribution function*  $\Phi(\cdot|x)$  of predictand  $W$  is defined by

$$\Phi(w|x) = \int_{-\infty}^w \phi(u|x) du. \quad (3)$$

It gives  $P(W \leq w|X = x) = \Phi(w|x)$ , where  $P$  stands for probability. The inverse function  $\Phi^{-1}(\cdot|x)$  is called the *posterior quantile function*. For any number  $p$ , such that  $0 < p < 1$ , and any deterministic forecast  $X = x$ , the *p-probability posterior quantile* of predictand  $W$  is the quantity  $w_p$  such that  $\Phi(w_p|x) = p$ . Therefrom,

$$w_p = \Phi^{-1}(p|x). \quad (4)$$

Equations (1)–(4) define the theoretic structure of the BPF. Equations (2)–(4) specify the three outputs, each of which constitutes the probabilistic forecast of  $W$ , given deterministic forecast  $x$ .

## 2.4 Bayesian Meta-Gaussian Model

The theoretic structure of the BPF can be implemented in many ways, as different mathematical models for  $g$  and  $f$  can be formulated, and different solution techniques for  $\phi$ ,  $\Phi$ , and  $\Phi^{-1}$  can be developed. A particularly elegant BPF is Gaussian-linear (Krzysztofowicz, 1983, 1987; Krzysztofowicz and Watada, 1986; Krzysztofowicz and Reese, 1991), but its applicability in meteorology is limited to a few predictands (with near-Gaussian  $g$ ) and a few short lead times (situations wherein a linear and homoscedastic dependence structure between  $X$  and  $W$  is mostly found).

Our objective is to propose a BPF of wide applicability — in that the predictand  $W$  and the predictor  $X$  are allowed to have distribution functions of any form, and the dependence structure between  $X$  and  $W$  is allowed to be nonlinear and heteroscedastic (which is the case with most meteorological forecasts, especially for longer lead times). This is the meta-Gaussian BPF (Kelly and Krzysztofowicz, 1995; Krzysztofowicz and Kelly, 2000a). In addition to the aforementioned advantages, it subsumes the Gaussian-linear BPF, and thus is a proper generalization thereof.

## 2.5 Modeling and Estimation

The remaining sections describe the modeling process, the estimation procedure, the goodness of fit to data, the statistical properties, and the practical advantages of the meta-Gaussian BPF. In all illustrations, the predictand is the daily maximum temperature; the forecast lead times are 24 h (1 day), 96 h (4 days), and 168 h (7 days) after 0000 UTC; the forecast points are the three stations: Savannah, Georgia (KSAV); Portland, Maine (KPWM); and Kalispell, Montana (KFCA).

### 3. THE PRIOR DISTRIBUTION FUNCTION

#### 3.1 Climatic Sample

Let  $W_k$  denote the maximum temperature on day  $k$  of the year ( $k = 1, \dots, 365$ ) at a given station. To keep the number of days in each year constant, for simplicity, 29 February is excluded. Because the time series  $\{W_k : k = 1, \dots, 365\}$  is obviously nonstationary, the climatic sample is formed for each day  $k$ . To increase the sample size, data are pooled from the consecutive five days centered on the given day. (The five-day window offers a reasonable compromise between increasing the sample size and precluding the nonstationarity effects.) Because of missing data, the sample size varies from day to day. To ensure uniformity, for comparison and convenience, the oldest data are removed from each day until the sample size for that day equals the smallest size among the 365 samples.

In summary (Table 1), for each day  $k$  of the year ( $k = 1, \dots, 365$ ), there is a climatic sample  $\{w_k(n) : n = 1, \dots, M\}$  of size  $M$ , where  $w_k(n)$  denotes the  $n$ th observation of the maximum temperature in the 5-day *sampling window* for day  $k$ .

#### 3.2 Standardization

The daily time series of sample deciles, the lowest observation, and the highest observation at Savannah are plotted in Fig. 2a. As expected, these time series confirm the nonstationarity of the maximum temperature. Hence, each predictand  $W_k$  has a different prior distribution function  $G_k$  ( $k = 1, \dots, 365$ ). Whereas modeling and estimation of 365 different distribution functions is feasible — and has been done by the NWS (Peroutka et al., 2005) — we seek a more efficient method of handling the nonstationarity.

The nonstationarity of the first two moments of  $W_k$  can be removed via standardization. First, for each day  $k$  ( $k = 1, \dots, 365$ ), the mean  $E(W_k) = m_k$  and the variance  $Var(W_k) = s_k^2$  are

estimated from the *original climatic sample*  $\{w_k(n) : n = 1, \dots, M\}$ . For operational forecasting, the time series of the estimates can be smoothed and approximated by first-order Fourier series expansions, as shown in Fig. 3; herein, they are used directly. Second, each observation is standardized:

$$w'_k(n) = \frac{w_k(n) - m_k}{s_k}, \quad n = 1, \dots, M. \quad (5)$$

The resultant *standardized climatic sample*  $\{w'_k(n) : n = 1, \dots, M\}$  is reanalyzed. Figure 2b shows the daily time series of the sample statistics: all decile time series are fairly flat, suggesting that (i) the seasonality of the maximum temperature is explained almost entirely by the seasonality of the mean and variance, and (ii) the prior distribution function  $G'_k$  of the standardized maximum temperature

$$W'_k = \frac{W_k - m_k}{s_k}, \quad k = 1, \dots, 365, \quad (6)$$

is approximately stationary. Of course, the standardization guarantees the stationarity of the first two moments:  $E(W'_k) = 0$  and  $Var(W'_k) = 1$  for  $k = 1, \dots, 365$ . The qualification “approximately” stationary is made, at least tentatively, because the time series of the lowest and the highest standardized observations exhibit some variability and a slight trend in the range, which appears wider in the warm season than in the cool season; however, these are extreme observations, and only two per day. Thus, their practical significance cannot be ascertained until the distribution functions are estimated.

### 3.3 Empirical Distribution Function

Results are reported for four diverse days in order to make a convincing case for the stationarity of the standardized maximum temperature. The chosen days are (Table 2): one of the coldest, one of the warmest, one with the maximum range in Fig. 2b, and one with the minimum range in

Fig. 2b.

From the standardized climatic sample for day  $k$ , the empirical distribution function of  $W'_k$  was constructed; it is specified by the set of  $M$  points

$$\{(w'_{k(n)}, p_n) : n = 1, \dots, M\}, \quad (7)$$

such that

$$P(W'_k \leq w'_{k(n)}) = p_n, \quad n = 1, \dots, M, \quad (8)$$

where  $w'_{k(n)}$  is the  $n$ th realization in the sample for day  $k$  sorted in the ascending order,  $p_n = n/(M + 1)$  is the Weibull plotting position, and  $M$  is the sample size.

### 3.4 Parametric Distribution Function

A parametric distribution function  $G'_k$  of  $W'_k$  was estimated and its goodness-of-fit to the empirical distribution function was evaluated in terms of the maximum absolute difference (MAD), a measure consistent with the Kolmogorov-Smirnov statistic:

$$MAD = \max_{1 \leq n \leq M} |p_n - G'_k(w'_{k(n)})|. \quad (9)$$

When there were several identical realizations in the sample, they formed a step in the empirical distribution function; the median plotting position in this step was used to calculate the absolute difference.

Excellent fits were obtained (Table 3) for all four days with  $G'_k$  coming from the Weibull family of distributions (Appendix A). Moreover, the parameter estimates were similar across the four days, implying that the standardization successfully removed the seasonality. Therefore, instead of estimating a different distribution function for every day of the year, it may be sufficient to estimate just one distribution function that will be valid for every day.

### 3.5 Stationary Prior Distribution Function

Under the stationarity hypothesis,  $W'_k = W'$  and  $G'_k = G'$  for  $k = 1, \dots, 365$ , where  $W'$  is the stationary-standardized maximum temperature on any day of the year, and  $G'$  is the stationary prior distribution function of  $W'$ .

To obtain  $G'$ , the standardized climatic samples for the four diverse days are pooled together and a single parametric distribution function is estimated from the pooled sample (Fig. 4). It is a Weibull distribution (Appendix A) with parameter estimates  $\alpha = 5.409$ ,  $\beta = 5.570$ , and  $\eta = -5$ . (Although temperature is not naturally bounded, a lower bound of  $-5$  barely restricts the sample space: under the standard normal distribution,  $P(W' < -5) = 0.2867 \times 10^{-6}$ .) The goodness-of-fit of this stationary parametric distribution function to the empirical distribution function for each day is reported in Table 3. When the stationary parameter estimates are used, the MAD is somewhat higher for two days and slightly lower, by chance, for the other two days. Overall, the fits remain very good ( $\text{MAD} < 0.05$ ), corroborating the stationarity hypothesis.

### 3.6 De-standardization

On any day  $k$  ( $k = 1, \dots, 365$ ), the transformation between the stationary distribution function  $G'$  of  $W'$  and the distribution  $G_k$  of  $W_k$  is

$$G_k(w) = G'\left(\frac{w - m_k}{s_k}\right), \quad (10)$$

giving  $P(W_k \leq w) = G_k(w)$  at any point  $w$  in the original sample space of the maximum temperature. The transformation between the corresponding density functions is

$$g_k(w) = \frac{1}{s_k} g'\left(\frac{w - m_k}{s_k}\right). \quad (11)$$

When  $G'$  is the Weibull distribution with parameters  $(\alpha, \beta, \eta)$ , it can be shown, via (10), that

$G_k$  is also the Weibull distribution with parameters  $(\alpha_k, \beta, \eta_k)$ , where

$$\begin{aligned}\alpha_k &= \alpha s_k, \\ \beta &= \beta, \\ \eta_k &= \eta s_k + m_k.\end{aligned}\tag{12}$$

Thus, the scale parameter  $\alpha_k$  and the shift parameter  $\eta_k$  capture the seasonality of the maximum temperature, whereas the shape parameter  $\beta$  captures the intrinsic, season-invariant, stochasticity of the maximum temperature.

The Weibull distribution functions  $G_k$  with the parameters calculated according to (12) are plotted along with the empirical distribution functions in the original sample spaces in Fig. 5. The nonstationarity of  $G_k$  is vivid, yet the fit remains excellent for each day. (The MAD defined in (9) remains invariant under the de-standardization.)

The above procedure, illustrated herein for Savannah, performed equally well for Portland and Kalispell, with the only distinction being that the best-fit parametric distribution for Portland turned out to be log-logistic rather than Weibull.

In conclusion, the prior distribution functions  $G_k$  for all days of the year ( $k = 1, \dots, 365$ ) can be specified by (i) a single, stationary, 3-parameter distribution (drawn from one of the common families), with parameters  $(\alpha, \beta, \eta)$  estimated from the pooled standardized sample, and (ii) the means and the standard deviations of the maximum temperatures  $\{(m_k, s_k) : k = 1, \dots, 365\}$  estimated from the climatic samples for all days. Once estimated, the prior distribution functions remain valid until enough additional climatic observations are collected to detect a change in climate since the last estimation.

### 3.7 Comparison of Models

The NWS took a different approach to modeling the nonstationary distribution functions  $G_k$  for  $k = 1, \dots, 365$  (Peroutka et al., 2005). It chose a four-parameter Generalized Lambda Distribution (GLD) of Karian and Dudewicz (2000); then it modeled the four time series of daily parameter estimates by employing cosine series which required estimating 22 constants. We calculated the parameter values for each of the four test days and evaluated the goodness-of-fit of the GLD to the empirical distribution function (Table 3). The fit of the GLD is decisively inferior to the fit of the Weibull distribution. In an extended test, we compared the two models on the first day of every month and then calculated the average MAD; it is 0.0270 for the Weibull model and 0.0456 for the GLD.

Overall, the model employing the standardization of the daily variates (as a means of obtaining a stationary time series) and a single, stationary, 3-parameter Weibull distribution fits the data better than the model employing a nonstationary four-parameter GLD. Moreover, the estimation of the nonstationary GLD is more complex because of the cosine series involved, and the use of the GLD is computationally far more demanding because there are no closed-form expressions for the distribution function and the density function. (The GLD is defined by its quantile function — the inverse of the distribution function.)



## 4. THE LIKELIHOOD FUNCTION

### 4.1 Joint Sample

Let  $x_k$  denote a deterministic forecast of the maximum temperature on day  $k$  of the year at a given station issued by a NWS field office with a specified lead time; the lead times considered herein are 24, 96, and 168 h after 0000 UTC. Let  $w_k$  denote the corresponding observation of the maximum temperature. The pair  $(x_k, w_k)$  forms a joint realization of  $(X_k, W_k)$ , where  $X_k$  is the forecast variate (the predictor, from the viewpoint of a rational decision maker), and  $W_k$  is the predictand. A joint sample  $\{(x_k, w_k)\}$  can be retrieved from the NDFD; however, its usage poses three challenges.

First, the available joint sample of  $(X_k, W_k)$  is typically much shorter than the climatic sample of  $W_k$ . For example, for the three stations analyzed herein, the joint samples are about one year long (Table 4) whereas the climatic samples are more than 100 years long (Table 1). The idea of augmenting the joint sample via simulation (Krzysztofowicz and Kelly, 2000a) or “re-forecasting” (Hamill et al., 2006) is not applicable here because the official NWS forecasts are subjective — they incorporate numerous human judgments made by different forecasters at different aggregation levels, or scales, as in national centers and field offices (Glahn and Ruth, 2003); and it is infeasible to determine, efficiently and reliably, the judgmental modifications of various model outputs that would have been made in years past by the currently employed NWS forecasters.

Second, the joint sample should be homogeneous: all forecasts included in it should have been produced by one forecast system — the same system for which the BPF is being developed. Inasmuch as the numerical weather prediction models, which provide guidance to forecasters, undergo modifications, and the forecasters in the NWS field offices change (as some relocate or retire and others are hired), the homogeneous joint sample may, in fact, be shorter than the sample

stored in the NDFD.

Third, the performance of the forecast system may be nonstationary; for instance, forecasts produced during a cool season may be more informative than forecasts produced during a warm season. This means that the stochastic dependence between  $X_k$  and  $W_k$ , which must be captured by the likelihood function, varies with season; in general, it varies with  $k$  ( $k = 1, \dots, 365$ ). Therefore, the likelihood function must be allowed to be nonstationary, despite the lack of an adequate sample to develop and estimate a formal nonstationary model.

## 4.2 Adaptive Scheme

To cope with the first challenge, it is necessary to formulate a statistical technique that can extract information from a small joint sample and then fuse it with information extracted from a large climatic sample. As this is the unique capability of the Bayesian approach, this approach becomes inevitable.

To cope with the other two challenges, an *adaptive scheme* is formulated for sampling, estimation, and forecasting as follows:

1. The joint sample is taken from about 90–120 days preceding a forecast day — the *sampling window* within which the system homogeneity and performance stationarity can be assumed. (The two time series  $\{X_k\}$  and  $\{W_k\}$  are still treated as nonstationary.)
2. The likelihood function estimated from a given joint sample is used in forecasting on each of the subsequent 5–10 days — the *forecasting window* to which the sampling assumptions are extended.

## 4.3 Joint Standardization

Like the time series of the predictand  $\{W_k : k = 1, \dots, 365\}$ , the time series of the forecast variate  $\{X_k : k = 1, \dots, 365\}$  is nonstationary. But unlike  $W_k$ , which could be characterized sta-

tistically using a large climatic sample,  $X_k$  cannot be characterized in the parallel manner because the joint sample is so small. Therefore, we resort to the following procedure.

First, each joint realization  $(x_k, w_k)$  is standardized using the climatic mean  $m_k$  and the climatic standard deviation  $s_k$  of predictand  $W_k$  for day  $k$ :

$$x'_k = \frac{x_k - m_k}{s_k}, \quad w'_k = \frac{w_k - m_k}{s_k}. \quad (13)$$

The plots of the original time series  $\{(x_k, w_k)\}$  and the standardized time series  $\{(x'_k, w'_k)\}$  in Fig. 6 for the lead time of 168 h show how this standardization eliminates, or at least reduces, the seasonality in the joint sample. Ditto for other lead times and stations. (Note how different the time series of the forecasts and observations are.)

Second, the standardized time series is assumed to be stationary and ergodic within a sampling window of  $N$  days. Realizations are retrieved from  $N$  days preceding the current forecast day (the day on which the forecast is to be made) to obtain a joint sample  $\{(x'(n), w'(n)) : n = 1, \dots, N\}$ . This is considered to be a random sample of the pair of variates  $(X', W')$  such that

$$X' = \frac{X_k - m_k}{s_k}, \quad W' = \frac{W_k - m_k}{s_k}, \quad (14)$$

for day  $k$  within the forecasting window for which a forecast is to be made with the specified lead time. While this standardization guarantees  $E(W') = 0$  and  $Var(W') = 1$ , no presumption is made, or needed, about the moments of  $X'$ .

The sampling windows and the sample sizes,  $N$ , chosen for the examples are listed in Table 4. Note that the joint samples include only past forecasts for which the corresponding observations are available on the forecast day. So, for the 168 h lead time, no forecasts issued within seven days of the forecast day are included. The sample sizes vary due to missing data.

#### 4.4 Marginal Distribution Function

Let  $\bar{K}'$  denote the marginal distribution function of the standardized forecast variate  $X'$ , such that  $\bar{K}'(x') = P(X' \leq x')$  for any  $x'$ . This  $\bar{K}'$  is to be estimated from the marginal sample  $\{x'(n) : n = 1, \dots, N\}$  of the standardized joint sample. (The overbar signifies that  $\bar{K}'$  is only an initial estimate of the marginal distribution function of  $X'$ ; this estimate can be revised later as a result of modeling (Krzysztofowicz and Kelly, 2000a, 2000b); but because the revised estimate is not needed for the operational BPF, it is not discussed herein.)

Figure 7 shows the empirical distribution functions constructed as explained in Section 3.3, and the estimated parametric distribution functions  $\bar{K}'$ , all Weibull, for two sampling windows (cool and warm) and two lead times (24 h and 168 h) at Savannah. The fits are good (Table 5). As the lead time increases, the distribution function becomes steeper and concentrated around the median. The effect of season is slight and is confounded by the difference in sample sizes.

A comparison of Fig. 7 with Fig. 4 reveals that in each of the four cases,  $\bar{K}' \neq G'$ ; most significantly, the tails of  $\bar{K}'$  are much shorter than the tails of  $G'$ . Clearly the forecasts have a distributional bias: it is evident in the shape parameter ( $\bar{\beta} > \beta$ ), but not in the scale parameter ( $\bar{\alpha} \simeq \alpha$ ), and it increases with lead time. Of course, the distributional bias implies a bias in both the mean and the variance of the forecast.

The general patterns shown here for Savannah hold also for Portland and Kalispell, with the exception of the distribution type, which is log-logistic at Portland.

The initial marginal distribution function  $\bar{K}_k$  of the forecast variate  $X_k$  on day  $k$ , with the specified lead time, is obtained from  $\bar{K}'$  via (14):

$$\bar{K}_k(x) = \bar{K}'\left(\frac{x - m_k}{s_k}\right), \quad (15)$$

giving  $P(X_k \leq x) = \bar{K}_k(x)$  at any point  $x$  in the original sample space of the maximum temperature. When the adaptive scheme (Section 4.2) is deployed,  $\bar{K}_k$  is nonstationary (because the climatic mean  $m_k$  and standard deviation  $s_k$  vary with day  $k$  of the year) even though the function  $\bar{K}'$  remains the same on all forecast days within the specified forecasting window. However,  $\bar{K}'$  is different for different lead times, as can be seen in Fig. 7.

When  $\bar{K}'$  is the Weibull distribution with parameters  $(\bar{\alpha}, \bar{\beta}, \bar{\eta})$ , it follows via (15) that  $\bar{K}_k$  is also the Weibull distribution with parameters  $(\bar{\alpha}_k, \bar{\beta}, \bar{\eta}_k)$ ; the relations between the corresponding parameters parallel those in (12). For example, the forecast of the maximum temperature at KSAV on 7 February made with the lead time of 168 h (7 days) is characterized by the variate  $X_{38}$  whose marginal distribution function  $\bar{K}_{38}$  is Weibull with parameters

$$\bar{\alpha}_{38} = \bar{\alpha} s_{38} = 5.341 \times 9.818 = 52.439,$$

$$\bar{\beta} = \bar{\beta} = 11.203,$$

$$\bar{\eta}_{38} = \bar{\eta} s_{38} + m_{38} = -5 \times 9.818 + 61.631 = 12.541,$$

where the values of  $\bar{\alpha}$ ,  $\bar{\beta}$ ,  $\bar{\eta}$  come from Fig. 7c and the values of  $m_{38}$ ,  $s_{38}$  come from Fig. 3. Four examples of  $\bar{K}_k$  for different days and lead times are shown in Fig. 8.

#### 4.5 Meta-Gaussian Likelihood Model

The objective now is to model the stochastic dependence between the standardized forecast variate  $X'$  and the standardized predictand  $W'$ . Towards this end, we employ the meta-Gaussian likelihood model of Krzysztofowicz and Kelly (2000a, 2000b). At the heart of this model is the *normal quantile transform* (NQT):

$$Z = Q^{-1}(\bar{K}'(X')), \quad V = Q^{-1}(G'(W')), \quad (16)$$

where  $Q$  is the standard normal distribution function and  $Q^{-1}$  is its inverse. The NQT guarantees

that in the sample space of the transformed variates  $(Z, V)$ , each marginal distribution function is Gaussian (Kelly and Krzysztofowicz, 1995, 1997), and several empirical studies using hydrological and meteorological data have demonstrated that the bivariate distribution function is Gaussian as well (e.g., Herr and Krzysztofowicz, 2005; Krzysztofowicz and Kelly, 2000a). Readers interested in these properties of the meta-Gaussian likelihood model are referred to the works cited above. The presentation below focuses on the practical estimation of the likelihood parameters.

#### 4.6 Likelihood Parameters

Given the standardized joint sample  $\{(x'(n), w'(n)) : n = 1, \dots, N\}$  for the specified lead time, the marginal distribution function  $\bar{K}'$  for the specified lead time, and the prior distribution function  $G'$ , each joint realization is processed through the NQT to obtain

$$z(n) = Q^{-1}(\bar{K}'(x'(n))), \quad v(n) = Q^{-1}(G'(w'(n))). \quad (17)$$

Then the transformed joint sample  $\{(z(n), v(n)) : n = 1, \dots, N\}$  is used to estimate the parameters  $a, b, \sigma$  of the Gaussian model:

$$E(Z|V = v) = av + b, \quad (18a)$$

$$Var(Z|V = v) = \sigma^2. \quad (18b)$$

The maximum likelihood estimators should be used (DeGroot, 1986).

Once its parameters are estimated, the Gaussian model (18) can be written in the form of the conditional quantile function:

$$z_{p|v} = av + b + \sigma Q^{-1}(p), \quad (19)$$

where  $0 < p < 1$ , and  $z_{p|v}$  is the  $p$ -probability quantile of  $Z$ , conditional on  $V = v$ ; that is  $P(Z \leq z_{p|v}|V = v) = p$ . As is well known, the linear regression equals the conditional median:  $E(Z|V = v) = z_{0.5|v}$ .

To visualize the dependence structure and to judge the fit of the model, scatterplots should

be produced, as shown in Fig. 9 for two lead times. The scatterplots of the transformed sample points  $(z(n), v(n))$  in Figs. 9c and 9d are linear and homoscedastic, and the Gaussian model (19), depicted by three conditional quantile functions, with  $p = 0.1, 0.5, 0.9$ , fits the sample points well. The scatterplots of the standardized sample points  $(x'(n), w'(n))$  in Figs. 9a and 9b are slightly nonlinear and heteroscedastic, especially for the 168-h lead time. The mapping of (19) into the standardized sample space gives

$$x'_{p|w'} = \bar{K}'^{-1}(Q(aQ^{-1}(G'(w')) + b + \sigma Q^{-1}(p))), \quad (20)$$

where  $x'_{p|w'}$  is the  $p$ -probability quantile of  $X'$ , conditional on  $W' = w'$ . In particular, the regression  $E(Z|V = v) = z_{0.5|v}$ , which is linear (Figs. 9c, 9d) is mapped into the conditional median  $x'_{0.5|w'}$ , which is slightly nonlinear (Figs. 9a, 9b). The 80% central credible interval about the conditional median of  $Z$ , whose width  $z_{0.9|v} - z_{0.1|v}$  is constant with  $v$  (Figs. 9c, d), is mapped into the 80% central credible interval about the conditional median of  $X'$ , whose width  $x'_{0.9|w'} - x'_{0.1|w'}$  decreases with  $w'$  (Figs. 9a, b).

In summary, these analyses demonstrate that applying the Gaussian model directly in the standardized sample space of  $X'$  and  $W'$  would be a wrong approach for the data at hand. Further evidence to this effect is given in Section 6.1.

#### 4.7 Forecast Informativeness

The stochastic dependence between  $Z$  and  $V$  is fully characterized by the Pearson's product-moment correlation coefficient  $\gamma$ , which may be expressed in terms of the likelihood parameters (Krzysztofowicz, 1992):

$$\gamma = (\text{sign of } a) \left[ \left( \frac{a}{\sigma} \right)^{-2} + 1 \right]^{-\frac{1}{2}}. \quad (21)$$

Under the meta-Gaussian likelihood model,  $\gamma$  remains a fully efficient measure of stochastic dependence between the standardized variates  $X'$  and  $W'$ , as well as between the original variates

$X$  and  $W$  (Kelly and Krzysztofowicz, 1997); and it can be transformed into the Spearman's rank correlation coefficient:

$$\rho = \frac{6}{\pi} \arcsin \left( \frac{\gamma}{2} \right). \quad (22)$$

The values of both measures,  $\gamma$  and  $\rho$ , are reported in Fig. 9.

Moreover, the Bayesian measure of informativeness of predictor  $X$  with respect to predictand  $W$  is the *informativeness score* (Krzysztofowicz, 1987, 1992):

$$IS = |\gamma| = \left[ \left( \frac{a}{\sigma} \right)^{-2} + 1 \right]^{-\frac{1}{2}}. \quad (23)$$

The score is bounded,  $0 \leq IS \leq 1$ , with  $IS = 0$  for an uninformative predictor, and  $IS = 1$  for a perfect predictor. (The informativeness score was called the Bayesian correlation score in the original publication by Krzysztofowicz (1992).) The value of  $IS$  is determined by the signal-to-noise ratio  $|a|/\sigma$ , estimated by regressing  $Z$  on  $V$ , as illustrated in Fig. 9: the absolute value of the slope coefficient,  $|a|$ , is the measure of signal, and the standard deviation of the residual,  $\sigma$ , is the measure of noise. Figure 9 shows that, as expected, the forecast with lead time of 24 h is significantly more informative than the forecast with lead time of 168 h.



## 5. PROBABILISTIC FORECAST

### 5.1 Posterior Parameters

After the likelihood parameters have been estimated, the posterior parameters can be calculated:

$$c_0 = \frac{-ab}{a^2 + \sigma^2}, \quad (24a)$$

$$c_1 = \frac{a}{a^2 + \sigma^2}, \quad (24b)$$

$$T^2 = \frac{\sigma^2}{a^2 + \sigma^2}. \quad (24c)$$

These parameters are for a specified lead time, and are valid for every day  $k$  within the forecasting window, as described in Section 4.2.

### 5.2 Forecasting Equations

To simplify the forecasting equations, the index  $k$  of the predictand's day is omitted, which is equivalent to the following substitutions in the forecasting problem. Given a deterministic forecast  $x = x_k$ , with a specified lead time, of predictand  $W = W_k$ , and given (i) the prior distribution function  $G = G_k$  and the corresponding prior density function  $g = g_k$  for day  $k$ , (ii) the marginal distribution function  $\bar{K} = \bar{K}_k$  for day  $k$  and for the specified lead time, and (iii) the posterior parameters  $c_0, c_1, T$  for the specified lead time, the probabilistic forecast is specified by one of the following constructs.

The posterior distribution function  $\Phi$  of  $W$ , defined in Eq. (3), is specified by the equation

$$\Phi(w) = Q \left( \frac{1}{T} [Q^{-1}(G(w)) - c_1 Q^{-1}(\bar{K}(x)) - c_0] \right), \quad (25)$$

where it is to be understood that  $\Phi(w) = \Phi(w|x)$ , with the abbreviated form being used when the value  $x$  of the deterministic forecast need not be shown. The  $p$ -probability posterior quantile of  $W$ , defined in Eq. (4), is specified by the equation

$$w_p = G^{-1} \left( Q(c_1 Q^{-1}(\bar{K}(x)) + c_0 + T Q^{-1}(p)) \right). \quad (26)$$

The posterior density function  $\phi$  of  $W$ , defined in Eq. (2), is specified by the equation

$$\phi(w) = \frac{1}{T} \exp \left( \frac{1}{2} \left\{ [Q^{-1}(G(w))]^2 - [Q^{-1}(\Phi(w))]^2 \right\} \right) g(w), \quad (27)$$

where it is understood that  $\phi(w) = \phi(w|x)$  when  $x$  need not be shown.

### 5.3 Posterior Functions

Figure 10 shows examples of probabilistic forecasts for Savannah, in the cool season, with the lead times of 24 h and 168 h. The posterior distribution functions (Figs. 10a, b), and the corresponding posterior density functions (Figs. 10c, d), quantify the uncertainty about the maximum temperature, given a deterministic forecast, either  $X = 50^\circ F$  or  $X = 70^\circ F$ . They also illustrate how the climatic uncertainty, quantified in terms of the prior distribution function (Figs. 10a, b) and the prior density function (Figs. 10c, d), is revised based on a deterministic forecast. Two effects of the forecast are apparent. First, the forecast shifts the center of the probability mass (under the density function) towards the forecasted temperature. Second, the forecast usually (but not always) reduces the uncertainty as the posterior density function is sharper than the prior density function; however, this reduction of uncertainty depends on the forecast and the lead time. Forecast  $X = 70^\circ F$  reduces the uncertainty more than forecast  $X = 50^\circ F$ , which reflects the heteroscedasticity of the dependence structure between  $X'$  and  $W'$  that was captured by the likelihood function. But as the lead time increases, from 24 h to 168 h, the degree by which the forecast reduces the climatic uncertainty diminishes: at 168 h, the posterior density functions are only slightly sharper than the prior density function.

## 5.4 Limit of Predictability

One of the most important properties of the BPF is its limiting behavior: when the informativeness score  $IS$  decreases to zero, the posterior distribution function  $\Phi$  converges uniformly to the prior distribution function  $G$ . (Likewise  $\phi$  converges to  $g$ .) Why is this practically important? The example (Fig. 9) illustrates a common property of meteorological forecasts: the longer the lead time, the lower the informativeness of the forecast. (This is further illustrated in Section 6.1.) At the limit of predictability, the forecast contains no signal ( $a = 0$ ), just noise ( $\sigma > 0$ ); thus  $IS = 0$ . It follows that as the lead time approaches the limit of predictability, the BPF guarantees that the posterior distribution functions  $\Phi$  converge to the prior distribution function  $G$  — the behavior that is evident already in Fig. 10. At the limit of predictability,  $\Phi = G$ , which is the climatic probabilistic forecast for the day because  $G$  was estimated from the climatic sample for the day. In conclusion, at the limit of predictability and beyond, the BPF automatically provides the decision maker with the correct and complete assessment of the climatic uncertainty.

## 6. DISCUSSION

### 6.1 Forms of Heteroscedasticity

The heteroscedasticity of the dependence structure between  $X'$  and  $W'$  may take many forms. That is why a flexible model for the likelihood function, such as the meta-Gaussian model, is necessary in order to correctly capture the forecast uncertainty. For instance, at Savannah, the conditional variance of  $X'$  decreases with  $w'$  (the 80% central credible interval in Fig. 9a narrows). But at Portland, the conditional variance of  $X'$  increases with  $w'$  (the 80% central credible interval in Fig. 11a widens). Loosely speaking, the 24-h forecasts are “better” for higher temperatures at Savannah, but for lower temperatures at Portland.

For the decision maker, the practical effect of this heteroscedasticity resurfaces in the posterior density function. As the forecasted temperature increases, the posterior density function becomes sharper at Savannah (Fig. 10c) — implying decreasing uncertainty, but grows flatter at Portland (Fig. 12c) — implying increasing uncertainty.

### 6.2 Revision of Uncertainty

The probabilistic forecasts for Portland (Fig. 12) illustrate also the interaction between the climatic uncertainty about the predictand and the informativeness of the forecast. The prior density functions (Figs. 12c, d) indicate that the climatic uncertainty is larger on the cool day than on the warm day. The likelihood parameters (Figs. 11c, d) indicate that the 24-h forecast is more informative on a cool day than on a warm day. The Bayesian revisions of (i) a larger climatic uncertainty based on a more informative forecast (for the cool day) and (ii) a smaller climatic uncertainty based on a less informative forecast (for the warm day) turn out to give nearly identical posterior uncertainty: when the plots in Fig. 12c are shifted  $2^\circ F$  to the right and then superposed on the plots in Fig. 12d, the posterior density functions match (almost). This illustrates the compensatory

nature of the interaction between the climatic uncertainty and the forecast informativeness, which only Bayes theorem can capture correctly and fully (in the shape of the revised density function).

### 6.3 Comparison of Performance

Figure 13 presents a compilation of the likelihood parameters  $a$ ,  $b$ , and  $\sigma$  for three stations, two seasons, and three lead times. The compilation includes also the informativeness score,  $IS$ , for the same stations, seasons, and seven lead times ranging from 24 h to 168 h. As the lead time increases, the signal measure  $a$  decreases, the noise measure  $\sigma$  increases (with one exception, explainable by the variability due to small sample size — see Table 4), and, consequently, the informativeness score  $IS$  decreases (again, with one exception). The informativeness score is the most stable measure. It declines at a similar rate for each station and season up to the lead time of 96 h. After that the decline is more rapid for all but one case, and the differences between the stations and the seasons grow larger. Finally, the bias parameter  $b$  is largely independent of the lead time, but highly variable across stations and seasons.

### 6.4 Generalization of Parameters

For implementation of the BPF in the NDFD, where it is desirable to store as few parameters as possible, an attempt should be made to determine which parameters vary least across stations, seasons, and lead times, and thus could be generalized. Figure 13 suggests the possibilities. Since each,  $IS$  and  $a$ , exhibits a relatively small variability across stations and seasons, perhaps a simple model could describe each of them as a function of the lead time. Then the values of  $IS$  and  $a$  could be used to calculate  $\sigma$  via Eq. (23). Therefore, only the two functions describing  $IS$  and  $a$  and the values of  $b$  would need to be stored.

## 7. CLOSURE

We have presented a theoretically-based, empirically-validated Bayesian processor that can be attached to the NDFD in order to process a deterministic forecast of a continuous predictand into a full-fledged probabilistic forecast. This probabilistic forecast is specified by a distribution function, or a density function, or a quantile function. Thus, it provides the complete characterization of uncertainty needed by rational decision makers who use formal decision models and by information providers who want to extract various forecast products for their customers.

The BPF offers the users a quintessential property: the posterior probability of any event can be taken at its face value. This is so because the Bayesian probabilistic forecast is well-calibrated against the climatic distribution function of the predictand, which is estimated from the longest available homogeneous sample that can be retrieved from the NCDC or some other archive.

The BPF comes furnished with a statistical procedure which copes with several non-stationarities of the stochastic processes involved and which ensures a parsimonious parameterization. In effect, all archived data pertinent to a given predictand and forecast lead time are processed into two samples for parameter estimation: the climatic sample of the predictand, and the joint sample of the predictor-predictand pair. The two samples may be of different sizes, as is almost always the case in meteorology. The unique attribute of the BPF is that it extracts information from each sample and then fuses it according to Bayes theorem, thereby ensuring that the resultant probabilistic forecast is always well calibrated and most informative, given the data at hand.

*Acknowledgments.* This material is based upon work supported by the National Science Foundation under Grant No. ATM-0135940, “New Statistical Techniques for Probabilistic Weather Forecasting”. The Meteorological Development Laboratory of the National Weather Service provided the data. Matthew R. Peroutka introduced the authors to the NDFD and the uncertainty quantification problem; his collaboration is much appreciated.

## APPENDIX A

### WEIBULL DISTRIBUTION

Let  $Y$  be a continuous variate having a bounded-below sample space  $(\eta, \infty)$ . Let  $\eta < y < \infty$  and  $0 < p < 1$ . Then variate  $Y$  has a Weibull distribution with the scale parameter  $\alpha > 0$ , the shape parameter  $\beta > 0$ , and the shift parameter  $-\infty < \eta < \infty$  if the distribution function of  $Y$  is

$$H(y) = 1 - \exp \left[ - \left( \frac{y - \eta}{\alpha} \right)^\beta \right], \quad (\text{A.1})$$

the density function of  $Y$  is

$$h(y) = \frac{\beta}{\alpha} \left( \frac{y - \eta}{\alpha} \right)^{\beta-1} \exp \left[ - \left( \frac{y - \eta}{\alpha} \right)^\beta \right], \quad (\text{A.2})$$

and the quantile function of  $Y$  is

$$H^{-1}(p) = \alpha [-\ln(1 - p)]^{\frac{1}{\beta}} + \eta. \quad (\text{A.3})$$

For properties and estimation methods, see Johnson and Kotz (1970, vol. 1), and Schütte et al. (1987).

## REFERENCES

- Alexandridis, M.G., and R. Krzysztofowicz, 1985: Decision models for categorical and probabilistic weather forecasts. *Applied Mathematics and Computation*, **17**, 241–266.
- DeGroot, M.H., 1986: *Probability and Statistics*, 2nd Edition, Addison-Wesley.
- Glahn, H.R., and D.P. Ruth, 2003: The new digital forecast database of the National Weather Service. *Bulletin of the American Meteorological Society*, **84**(2), 195–201.
- Hamill, T.M., J.S. Whitaker, and S.L. Mullen, 2006: Reforecasts: An important dataset for improving weather predictions. *Bulletin of the American Meteorological Society*, **87**(1), 33–46.
- Herr, H.D., and R. Krzysztofowicz, 2005: Generic probability distribution of rainfall in space: The bivariate model. *Journal of Hydrology*, **306**(1–4), 234–263.
- Johnson, N., and S. Kotz, 1970: *Distributions in Statistics: Continuous Univariate Distributions*, vol. 1 and 2, Wiley.
- Karian, Z.A., and E.J. Dudewicz, 2000: *Fitting Statistical Distributions: The Generalized Lambda Distribution and Generalized Bootstrap Methods*, CRC Press.
- Kelly, K.S., and R. Krzysztofowicz, 1995: Bayesian revision of an arbitrary prior density. *Proceedings of the Section on Bayesian Statistical Science*, American Statistical Association, 50–53.
- Kelly, K.S., and R. Krzysztofowicz, 1997: A bivariate meta-Gaussian density for use in hydrology. *Stochastic Hydrology and Hydraulics*, **11**(1), 17–31.
- Krzysztofowicz, R., 1983: Why should a forecaster and a decision maker use Bayes theorem. *Water Resources Research*, **19**(2), 327–336.



- Krzysztofowicz, R., 1987: Markovian forecast processes. *Journal of the American Statistical Association*, **82**(397), 31–37.
- Krzysztofowicz, R., 1992: Bayesian correlation score: A utilitarian measure of forecast skill. *Monthly Weather Review*, **120**(1), 208–219.
- Krzysztofowicz, R., 1999: Bayesian forecasting via deterministic model. *Risk Analysis*, **19**(4), 739–749.
- Krzysztofowicz, R., and K.S. Kelly, 2000a: Hydrologic uncertainty processor for probabilistic river stage forecasting. *Water Resources Research*, **36**(11), 3265–3277.
- Krzysztofowicz, R., and K.S. Kelly, 2000b: Bayesian improver of a distribution. *Stochastic Environmental Research and Risk Assessment*, **14**(6), 449–470.
- Krzysztofowicz, R., and S. Reese, 1991: Bayesian analyses of seasonal runoff forecasts. *Stochastic Hydrology and Hydraulics*, **5**(4), 295–322.
- Krzysztofowicz, R., and L.M. Watada, 1986: Stochastic model of seasonal runoff forecasts. *Water Resources Research*, **22**(3), 296–302.
- Peroutka, M.R., G.J. Zylstra, and J.L. Wagner, 2005: Assessing forecast uncertainty in the National Digital Forecast Database. *Preprints, 21st Conference on Weather Analysis and Forecasting*, Washington D.C., American Meteorological Society; paper number 2B.3.
- Ryan, R.T., 2003: Digital forecasts: Communication, public understanding, and decision making. *Bulletin of the American Meteorological Society*, **8**, 1001–1005.
- Schütte, T., O. Salka, and S. Israelsson, 1987: The use of the Weibull distribution for thunderstorm parameters. *Journal of Climate and Applied Meteorology*, **26**, 457–463.

Table 1. Climatic samples for estimation of the prior distribution functions.

Station	Start date	End date	Years	Days	$M$
KSAV	Jan. 1874	Dec. 2001	128	45,613	618
KPWM	Apr. 1874	Dec. 2001	128	46,654	634
KFCA	June 1896	Dec. 2001	106	38,457	519

Table 2. Days and climatic sample statistics for the prior distribution functions at KSAV.

Day's attribute	Date	$k$	$m_k[^\circ F]$	$s_k[^\circ F]$
cool	31 Jan.	31	61.743	9.894
warm	1 Aug.	213	90.686	4.452
max range	20 Sep.	263	84.427	5.717
min range	15 Nov.	319	69.764	7.829

Table 3. Maximum absolute difference (MAD) between the empirical prior distribution function of  $W'_k$  for the given day  $k$  and an estimated parametric prior distribution function at KSAV.

Distribution	31 Jan.	1 Aug.	20 Sep.	15 Nov.
Weibull for given day	0.0161	0.0242	0.0158	0.0333
Stationary Weibull	0.0158	0.0307	0.0387	0.0290
GLD for given day	0.0371	0.0396	0.0839	0.0827

Table 4. Joint samples for estimation of the likelihood functions.

Station	Available sample			Sampling window		Sample size $N$		
	Start date	End date	Years	Season	End date	24 h	96 h	168 h
KSAV	Oct. 2004	Feb. 2006	1.42	cool	31 Jan. 2005	116	113	110
				warm	1 Aug. 2005	66	63	62
KPWM	Oct. 2004	Feb. 2006	1.42	cool	31 Jan. 2005	115	112	109
				warm	1 Aug. 2005	67	62	60
KFCA	Oct. 2004	June 2005	0.75	cool	31 Jan. 2005	117	114	109
				warm	30 June 2005	74	71	70

Table 5. Maximum absolute difference (MAD) between the empirical marginal distribution function of  $X'$  and the estimated Weibull marginal distribution function of  $X'$  at KSAV.

Sampling window		Lead time		
Season	End date	24 h	96 h	168 h
cool	31 Jan. 2005	0.0372	0.0371	0.0376
warm	1 Aug. 2005	0.0736	0.0537	0.0500

## FIGURE CAPTIONS

- Figure 1. The concept of data fusion and uncertainty quantification via the Bayesian Processor of Forecast (BPF) coupled to the National Digital Forecast Database (NDFD) and the National Climatic Data Center (NCDC).
- Figure 2. Time series of the sample deciles, the lowest observation, and the highest observation of the daily maximum temperature at KSAV: (a) in the original sample space, and (b) in the standardized sample space.
- Figure 3. Time series of (a) the sample mean, and (b) the sample standard deviation of the daily maximum temperature at KSAV, and the fitted first-order Fourier series expansions.
- Figure 4. Stationary prior distribution functions, empirical and Weibull  $G'$ , of the standardized predictand  $W'$ , estimated from the pooled standardized climatic sample (for 31 Jan., 1 Aug., 20 Sep., 15 Nov.) at KSAV.
- Figure 5. Prior distribution functions, empirical (constructed from the climatic sample for day  $k$ ) and Weibull  $G_k$  (derived from the stationary prior distribution function  $G'$  shown in Fig. 4, the climatic mean  $m_k$ , and the climatic standard deviation  $s_k$  for day  $k$ ), in the original sample space of the predictand  $W_k$  at KSAV for: (a) 31 Jan., (b) 1 Aug., (c) 20 Sep., and (d) 15 Nov.
- Figure 6. Time series of the joint sample, forecasts with 168 h lead time and observations, at KSAV: (a) in the original sample space, and (b) in the standardized sample space.

Figure 7. Marginal distribution functions, empirical and Weibull  $\bar{K}'$ , of the standardized forecast variates  $X'$ , obtained using sampling windows for two forecast days and two lead times at KSAV: (a) cool day, 24 h lead time; (b) warm day, 24 h lead time; (c) cool day, 168 h lead time; (d) warm day, 168 h lead time.

Figure 8. Marginal distribution function  $\bar{K}_k$  of the forecast variate  $X_k$  (derived from the distribution function  $\bar{K}'$  shown in Fig. 7, the climatic mean  $m_k$ , and the climatic standard deviation  $s_k$  for day  $k$  shown in Fig. 3) at KSAV for: (a) cool day (1 Feb.), 24 h lead time; (b) warm day (2 Aug.), 24 h lead time; (c) cool day (7 Feb.), 168 h lead time; (d) warm day (8 Aug.), 168 h lead time.

Figure 9. Dependence structure of the likelihood function at KSAV obtained using the sampling window for a cool day (31 Jan.) and two lead times; meta-Gaussian regression of  $X'$  on  $W'$  and the 80% central credible interval in the standardized sample space: (a) 24 h lead time, (b) 168 h lead time; linear regression of  $Z$  on  $V$  and the 80% central credible interval in the normal sample space: (c) 24 h lead time, (d) 168 h lead time.

Figure 10. Quantification of uncertainty about the maximum temperature  $W$  at KSAV on a cool day. Prior distribution function  $G$  and posterior distribution functions  $\Phi$  conditional on deterministic forecasts  $X = x [^{\circ}F]$  issued with the lead times of (a) 24 h, for 1 Feb., (b) 168 h, for 7 Feb. Prior density function  $g$  and posterior density functions  $\phi$  conditional on deterministic forecasts  $X = x [^{\circ}F]$  issued with the lead times of (c) 24 h, for 1 Feb., (b) 168 h, for 7 Feb.



Figure 11. Dependence structure of the likelihood function for 24-h lead time at KPWM obtained using the sampling window for a cool day (31 Jan.) and a warm day (1 Aug.); meta-Gaussian regression of  $X'$  on  $W'$  and the 80% central credible interval in the standardized sample space: (a) cool day, (b) warm day; linear regression of  $Z$  on  $V$  and the 80% central credible interval in the normal sample space: (c) cool day, (d) warm day.

Figure 12. Quantification of uncertainty about the maximum temperature  $W$  at KPWM with 24-h lead time. Prior distribution function  $G$  and posterior distribution functions  $\Phi$  conditional on deterministic forecasts  $X = x$  [ $^{\circ}F$ ] for (a) cool day (1 Feb.), (b) warm day (2 Aug.). Prior density function  $g$  and posterior density functions  $\phi$  conditional on deterministic forecasts  $X = x$  [ $^{\circ}F$ ] for (c) cool day (1 Feb.), (d) warm day (2 Aug.).

Figure 13. Likelihood parameters at KSAV, KPWM, and KFCA for forecasting windows in the cool season (black symbols) and the warm season (open symbols), with the lead times of 24 h, 96 h, 168 h: (a) signal measure  $a$ , (b) bias measure  $b$ , (c) noise measure  $\sigma$ . The informativeness score  $IS$  for all three stations, two seasons, and seven lead times (d).

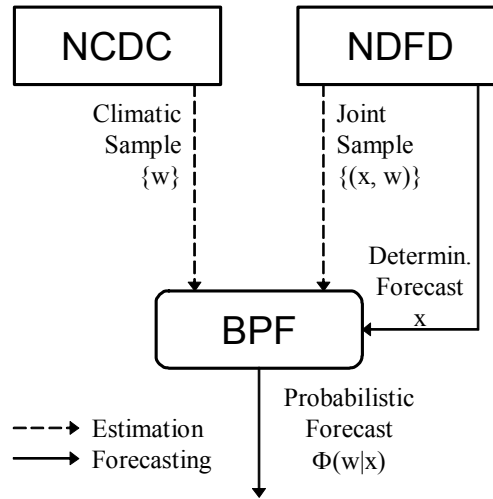


Figure 1. The concept of data fusion and uncertainty quantification via the Bayesian Processor of Forecast (BPF) coupled to the National Digital Forecast Database (NDFD) and the National Climatic Data Center (NCDC).

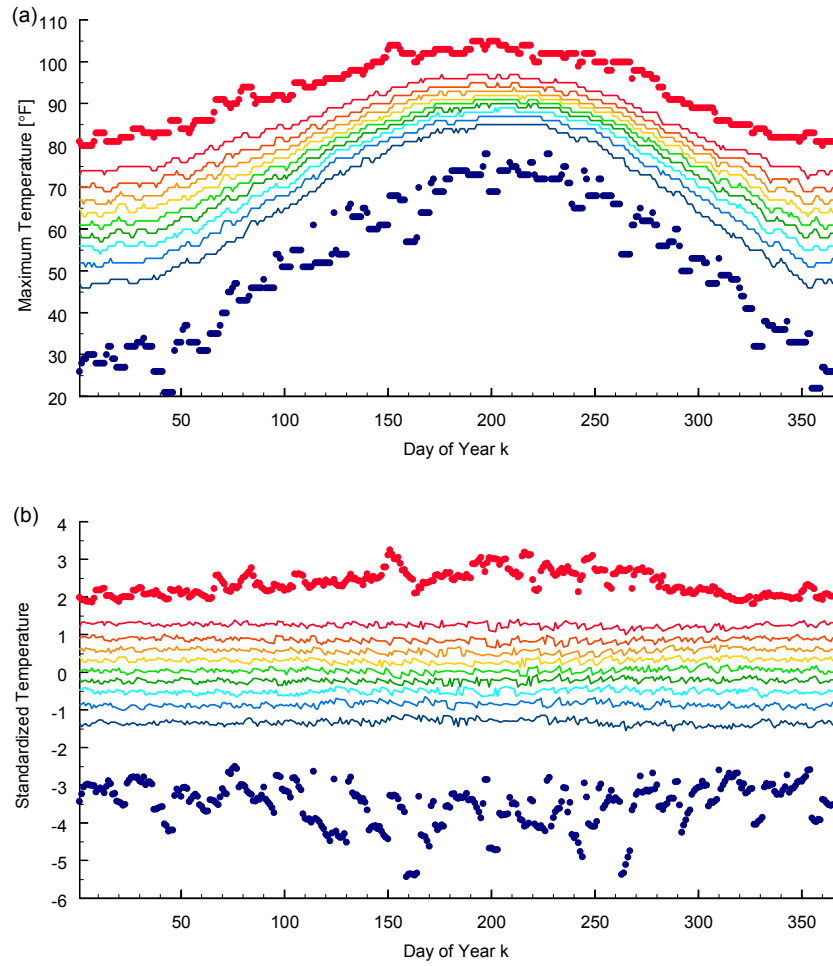


Figure 2. Time series of the sample deciles, the lowest observation, and the highest observation of the daily maximum temperature at KSAV: (a) in the original sample space, and (b) in the standardized sample space.

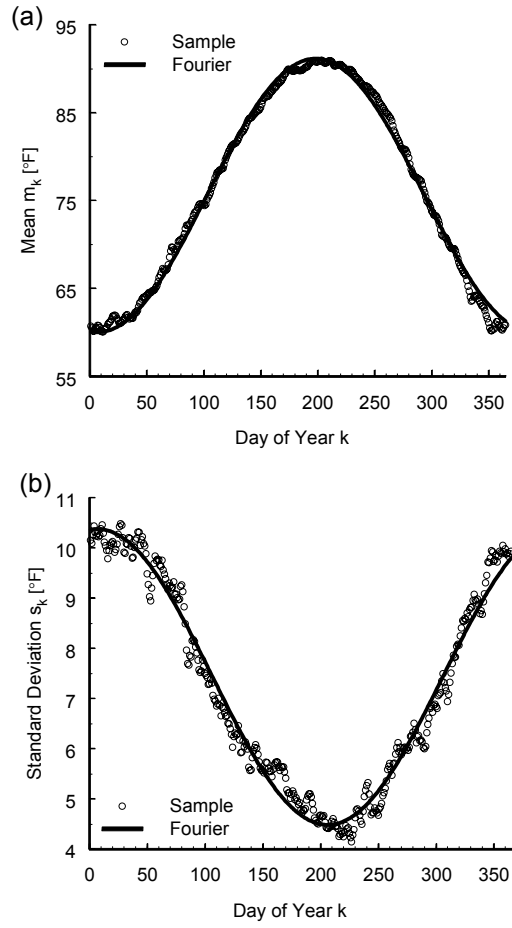


Figure 3. Time series of (a) the sample mean, and (b) the sample standard deviation of the daily maximum temperature at KSAV, and the fitted first-order Fourier series expansions.

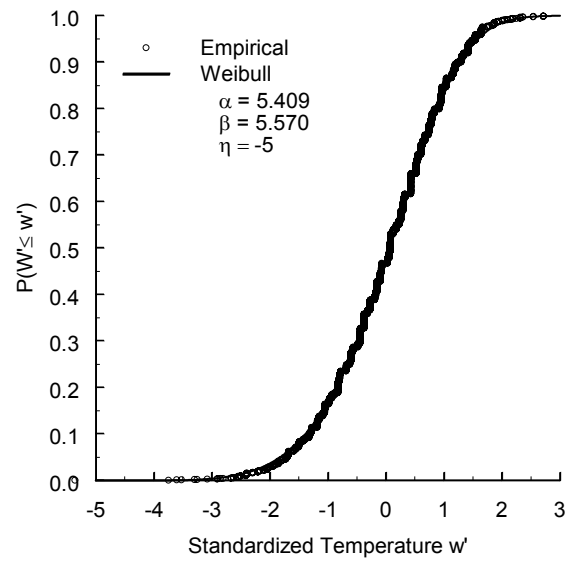


Figure 4. Stationary prior distribution functions, empirical and Weibull  $G'$ , of the standardized predictand  $W'$ , estimated from the pooled standardized climatic sample (for 31 Jan., 1 Aug., 20 Sep., 15 Nov.) at KSAV.

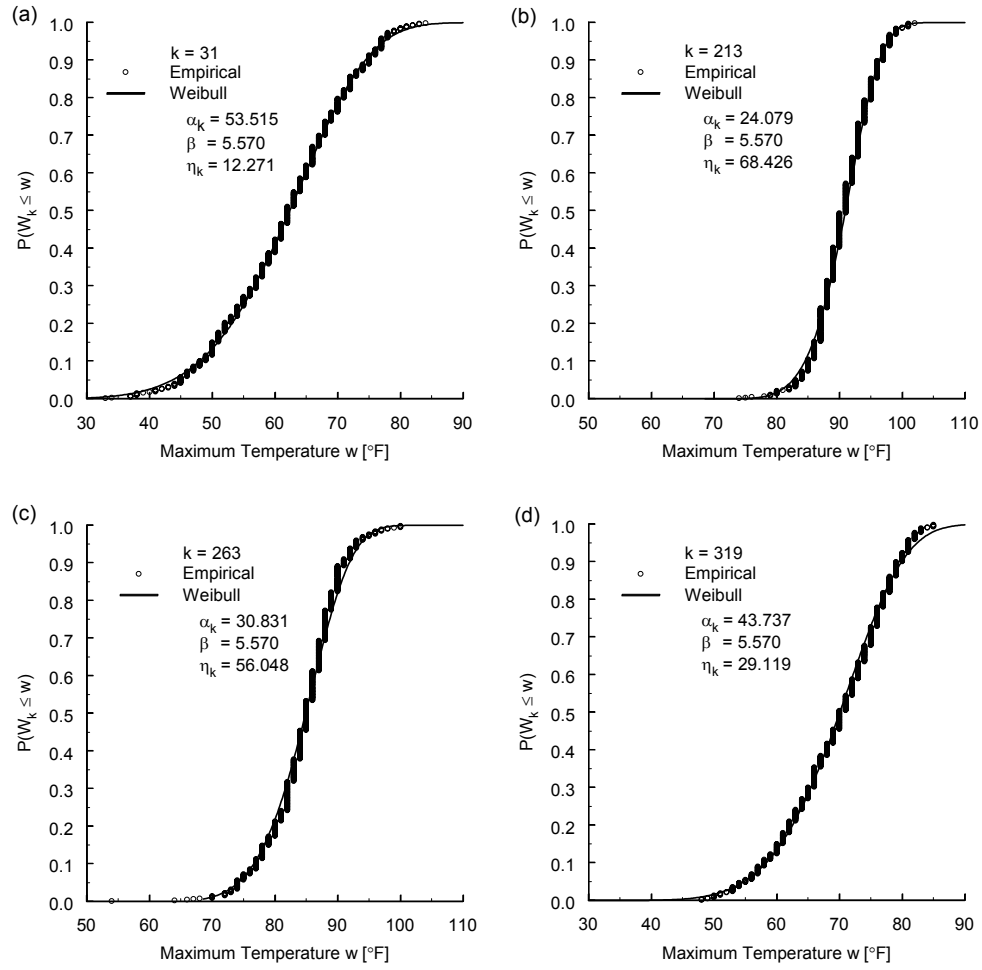


Figure 5. Prior distribution functions, empirical (constructed from the climatic sample for day  $k$ ) and Weibull  $G_k$  (derived from the stationary prior distribution function  $G'$  shown in Fig. 4, the climatic mean  $m_k$ , and the climatic standard deviation  $s_k$  for day  $k$ ), in the original sample space of the predictand  $W_k$  at KSAV for: (a) 31 Jan., (b) 1 Aug., (c) 20 Sep., and (d) 15 Nov.

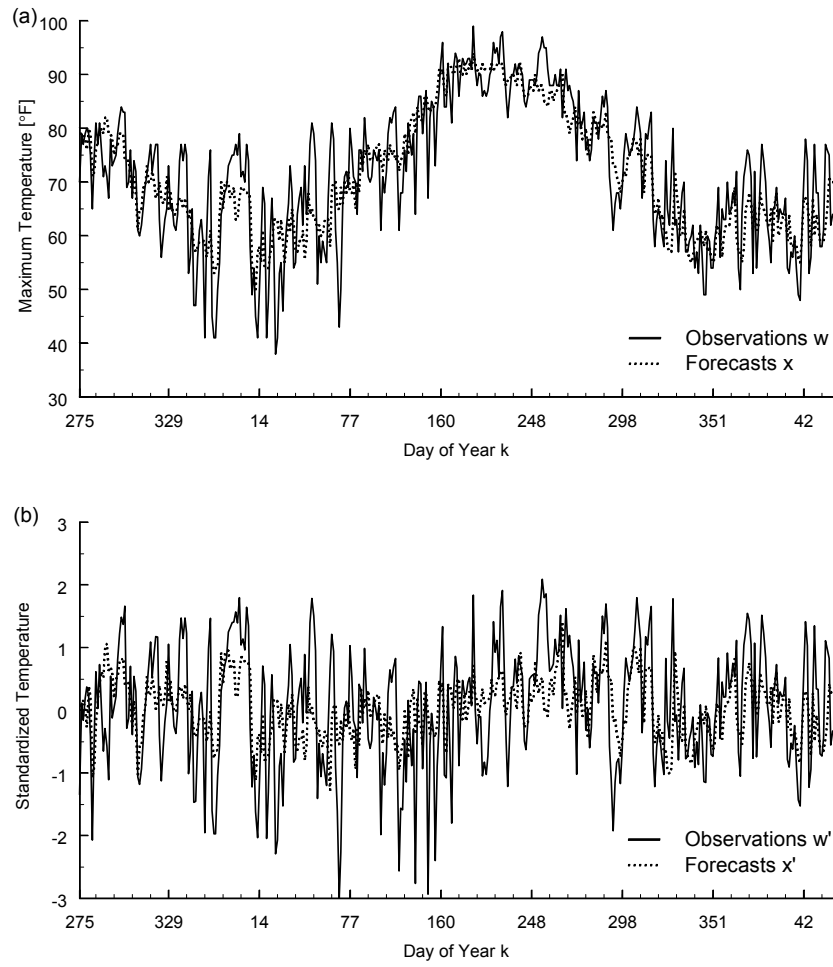


Figure 6. Time series of the joint sample, forecasts with 168 h lead time and observations, at KSAV: (a) in the original sample space, and (b) in the standardized sample space.

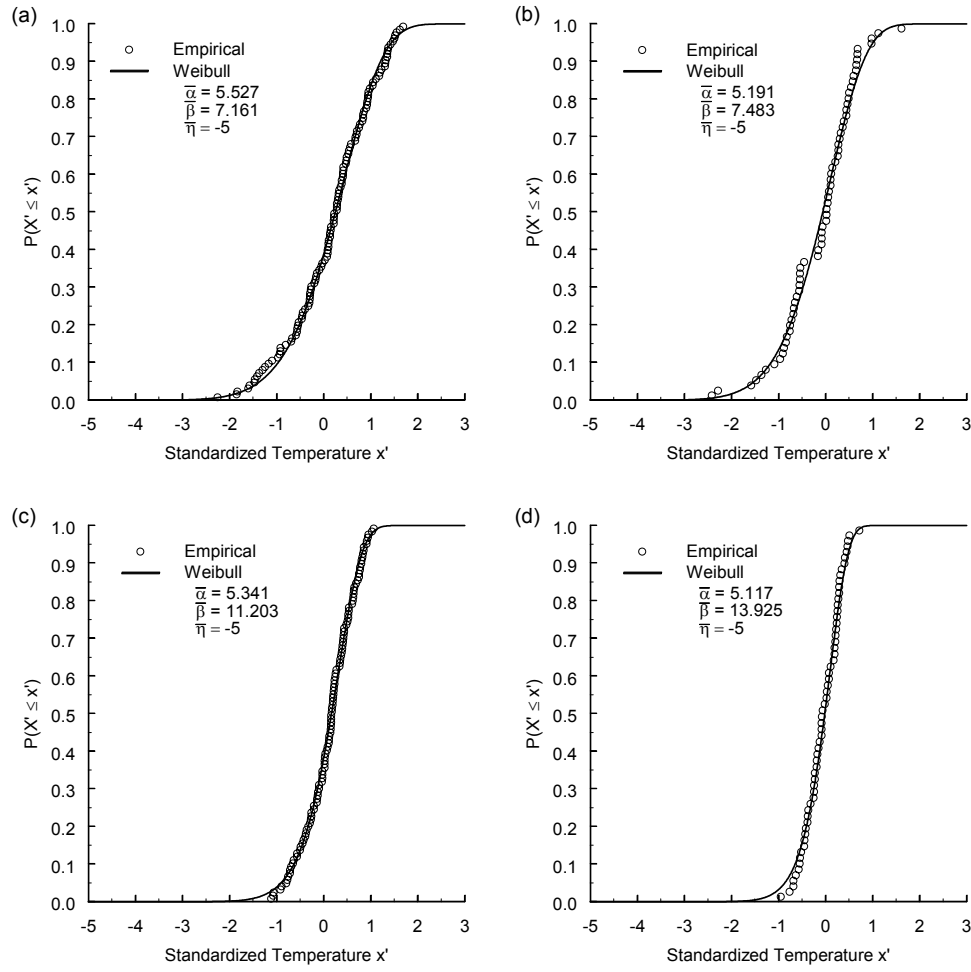


Figure 7. Marginal distribution functions, empirical and Weibull  $\bar{K}'$ , of the standardized forecast variates  $X'$ , obtained using sampling windows for two forecast days and two lead times at KSAV: (a) cool day, 24 h lead time; (b) warm day, 24 h lead time; (c) cool day, 168 h lead time; (d) warm day, 168 h lead time.



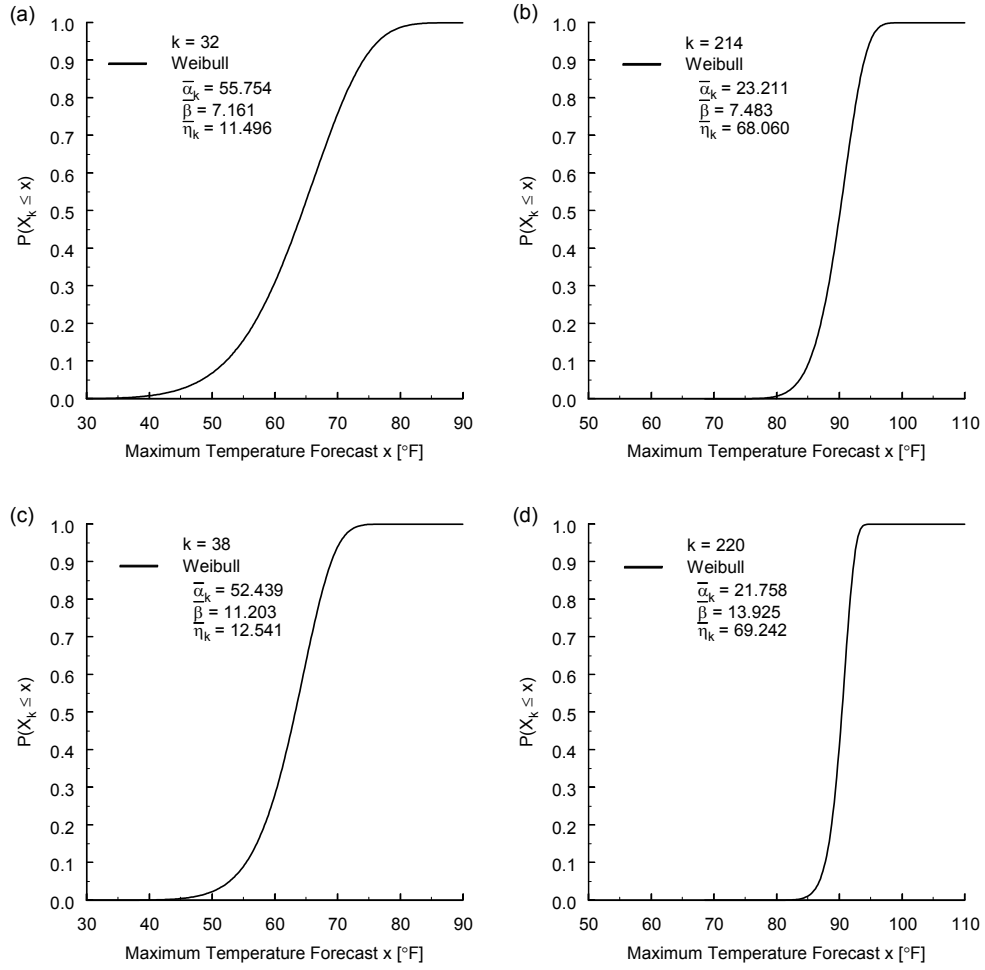


Figure 8. Marginal distribution function  $\bar{K}_k$  of the forecast variate  $X_k$  (derived from the distribution function  $\bar{K}'$  shown in Fig. 7, the climatic mean  $m_k$ , and the climatic standard deviation  $s_k$  for day  $k$  shown in Fig. 3) at KSAV for: (a) cool day (1 Feb.), 24 h lead time; (b) warm day (2 Aug.), 24 h lead time; (c) cool day (7 Feb.), 168 h lead time; (d) warm day (8 Aug.), 168 h lead time.

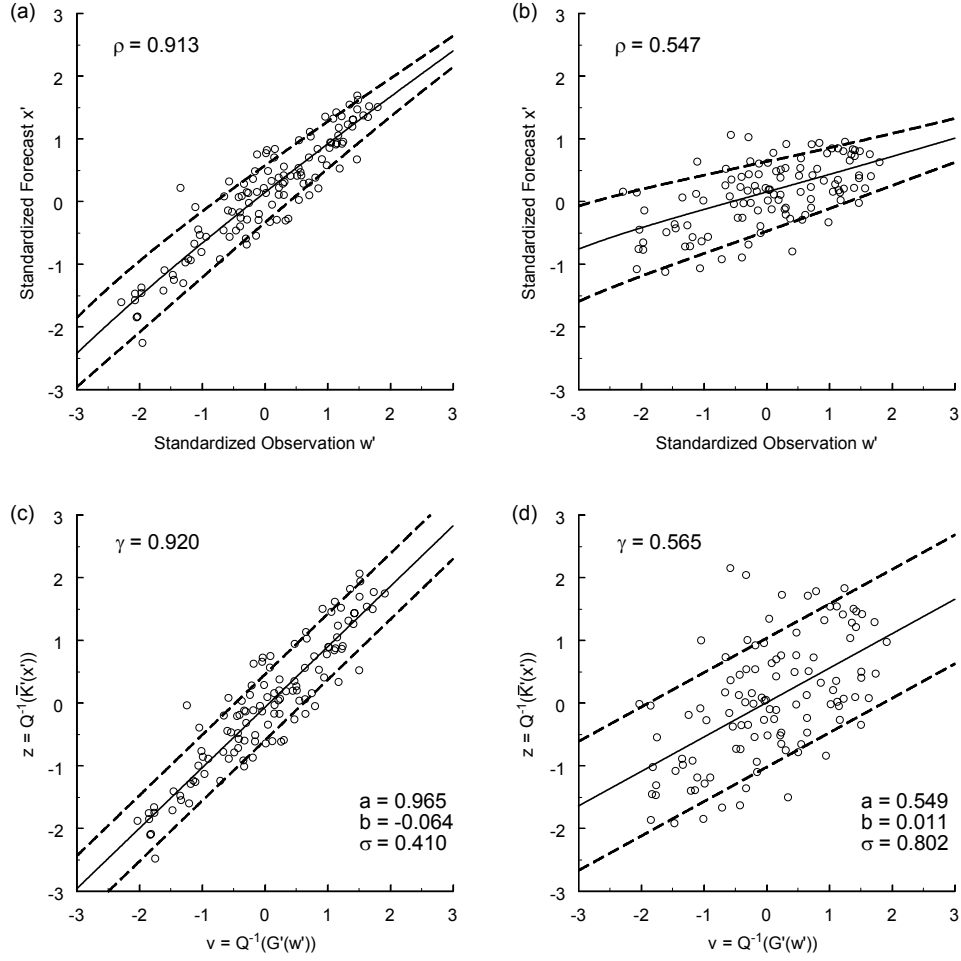


Figure 9. Dependence structure of the likelihood function at KSAV obtained using the sampling window for a cool day (31 Jan.) and two lead times; meta-Gaussian regression of  $X'$  on  $W'$  and the 80% central credible interval in the standardized sample space: (a) 24 h lead time, (b) 168 h lead time; linear regression of  $Z$  on  $V$  and the 80% central credible interval in the normal sample space: (c) 24 h lead time, (d) 168 h lead time.

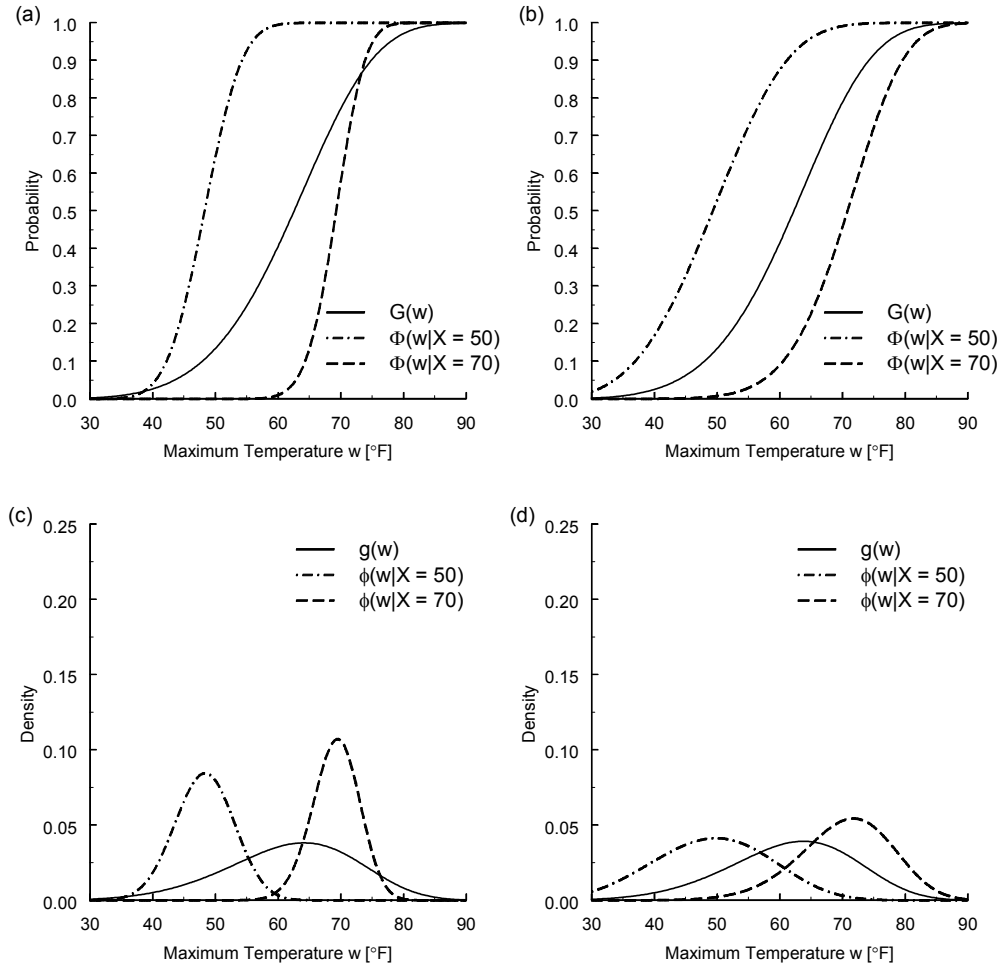


Figure 10. Quantification of uncertainty about the maximum temperature  $W$  at KSAV on a cool day. Prior distribution function  $G$  and posterior distribution functions  $\Phi$  conditional on deterministic forecasts  $X = x$  [°F] issued with the lead times of (a) 24 h, for 1 Feb., (b) 168 h, for 7 Feb. Prior density function  $g$  and posterior density functions  $\phi$  conditional on deterministic forecasts  $X = x$  [°F] issued with the lead times of (c) 24 h, for 1 Feb., (b) 168 h, for 7 Feb.

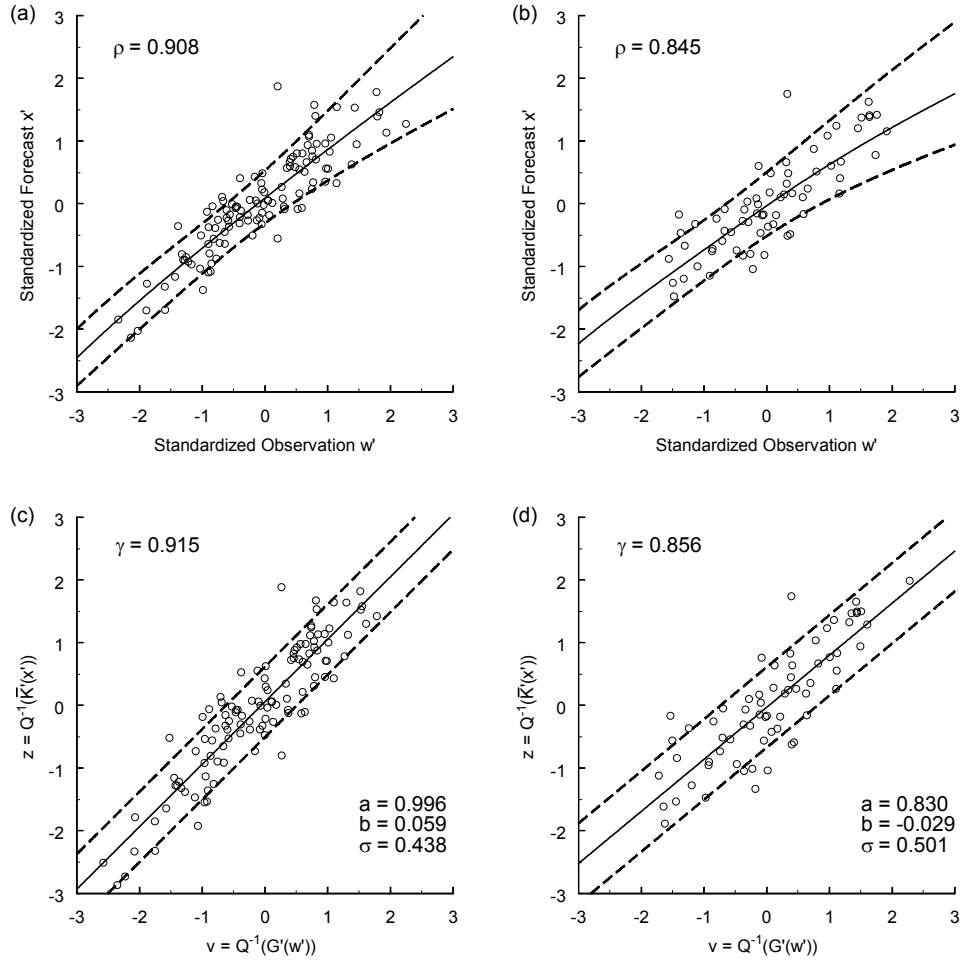


Figure 11. Dependence structure of the likelihood function for 24-h lead time at KPWM obtained using the sampling window for a cool day (31 Jan.) and a warm day (1 Aug.); meta-Gaussian regression of  $X'$  on  $W'$  and the 80% central credible interval in the standardized sample space: (a) cool day, (b) warm day; linear regression of  $Z$  on  $V$  and the 80% central credible interval in the normal sample space: (c) cool day, (d) warm day.

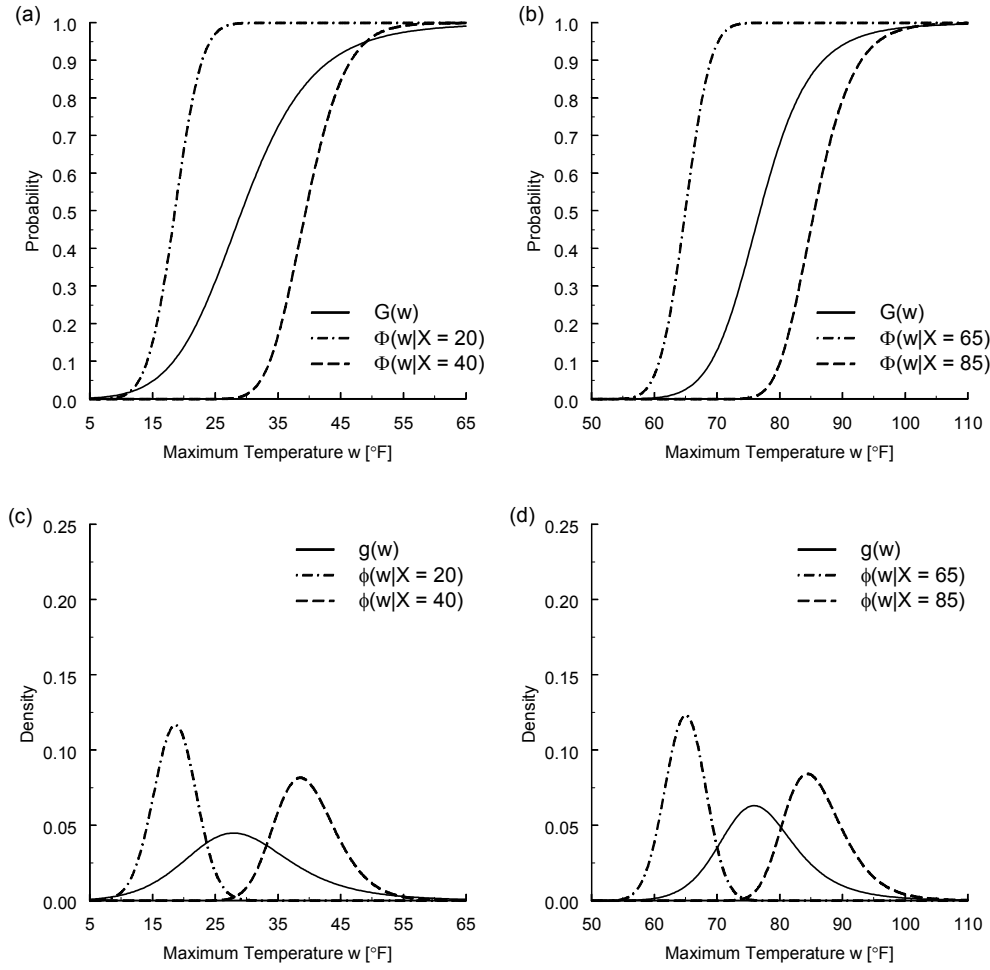


Figure 12. Quantification of uncertainty about the maximum temperature  $W$  at KPWM with 24-h lead time. Prior distribution function  $G$  and posterior distribution functions  $\Phi$  conditional on deterministic forecasts  $X = x$  [ $^{\circ}F$ ] for (a) cool day (1 Feb.), (b) warm day (2 Aug.). Prior density function  $g$  and posterior density functions  $\phi$  conditional on deterministic forecasts  $X = x$  [ $^{\circ}F$ ] for (c) cool day (1 Feb.), (d) warm day (2 Aug.).

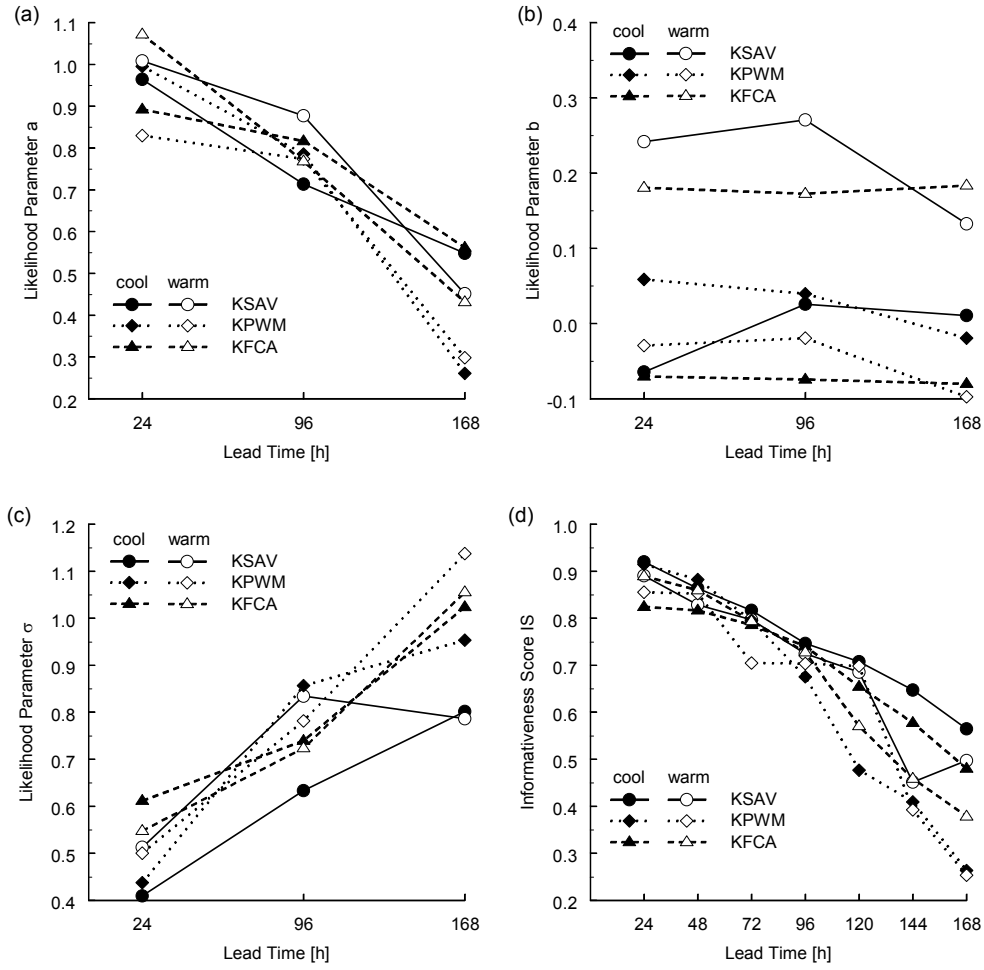


Figure 13. Likelihood parameters at KSAV, KPWM, and KFCA for forecasting windows in the cool season (black symbols) and the warm season (open symbols), with the lead times of 24 h, 96 h, 168 h: (a) signal measure  $a$ , (b) bias measure  $b$ , (c) noise measure  $\sigma$ . The informativeness score  $IS$  for all three stations, two seasons, and seven lead times (d).